

# *Análisis de datos provenientes de Redes Sociales*

Rolando Vicuña M.



**VicData**

“La revolución de los datos”





“La revolución de los datos”

# ¿Quiénes Somos?

- **VicData es una empresa Ecuatoriana creada para satisfacer las nuevas demandas de mercado en lo referente al aprovechamiento de datos, por esto su estrategia se basa en Investigación + Desarrollo + Innovación de herramientas y metodologías que permitan sacar el máximo partido de los datos a gran escala, para una toma de decisiones oportuna y acertada.**



“La revolución de los datos”

# Aprovechamiento de datos Provenientes de Twitter

- De acuerdo a un informe sobre indicadores de desarrollo mundial realizado en el año 2015 por el Banco Mundial, se evidenció que en un día típico en el uso del Internet se generan 803 millones de Tuits.
- Twitter pone a disposición toda esta información en tiempo real mediante un API de uso público.
- Actualmente se están desarrollando diferentes investigaciones con el consumo de esta información para conocer lo que la gente siente respecto a una temática que se este tuiteando.
- La información de redes sociales es posible aprovecharla, siempre que la misma pase por un proceso de limpieza y validación, para detectar información producida por robots y descartar información poco útil.



“La revolución de los datos”

# Demostración del aprovechamiento de datos provenientes de Twitter

- **Ejercicio1:**

¿Cómo el Big Data puede apoyar en la Gestión en Desastres Naturales?, con aplicación a los sismos producidos este año en Ecuador.

- Captura de Tuits producidos durante el sismo.
- Almacenamiento y depuración de Tuits.
- Aplicación de Análisis de Sentimientos utilizando Machine Learning.
- La información captada principalmente fue: Usuario, mensaje y georreferenciación.

- **Ejercicio2:**

Evaluación del sentir de la población respecto a temas de interés nacional.

- Se analizaron los temas mas tuiteados.
- Se realizó un análisis utilizando series temporales para determinar el cambio de sentimientos de la población respecto a estas temáticas.



“La revolución de los datos”

# ¿Herramientas Utilizadas?

- **Captura:**

El proceso de captura de datos lo realizamos mediante Spark Streaming y utilizando el API de Twitter.

- **Almacenamiento:**

Para almacenar la información utilizamos ElasticSearch para los datos calientes y para los datos fríos Hadoop.

- **Procesamiento:**

El procesamiento lo realizamos con Spark para la aplicación de técnicas de Machine Learning.

- **Análisis:**

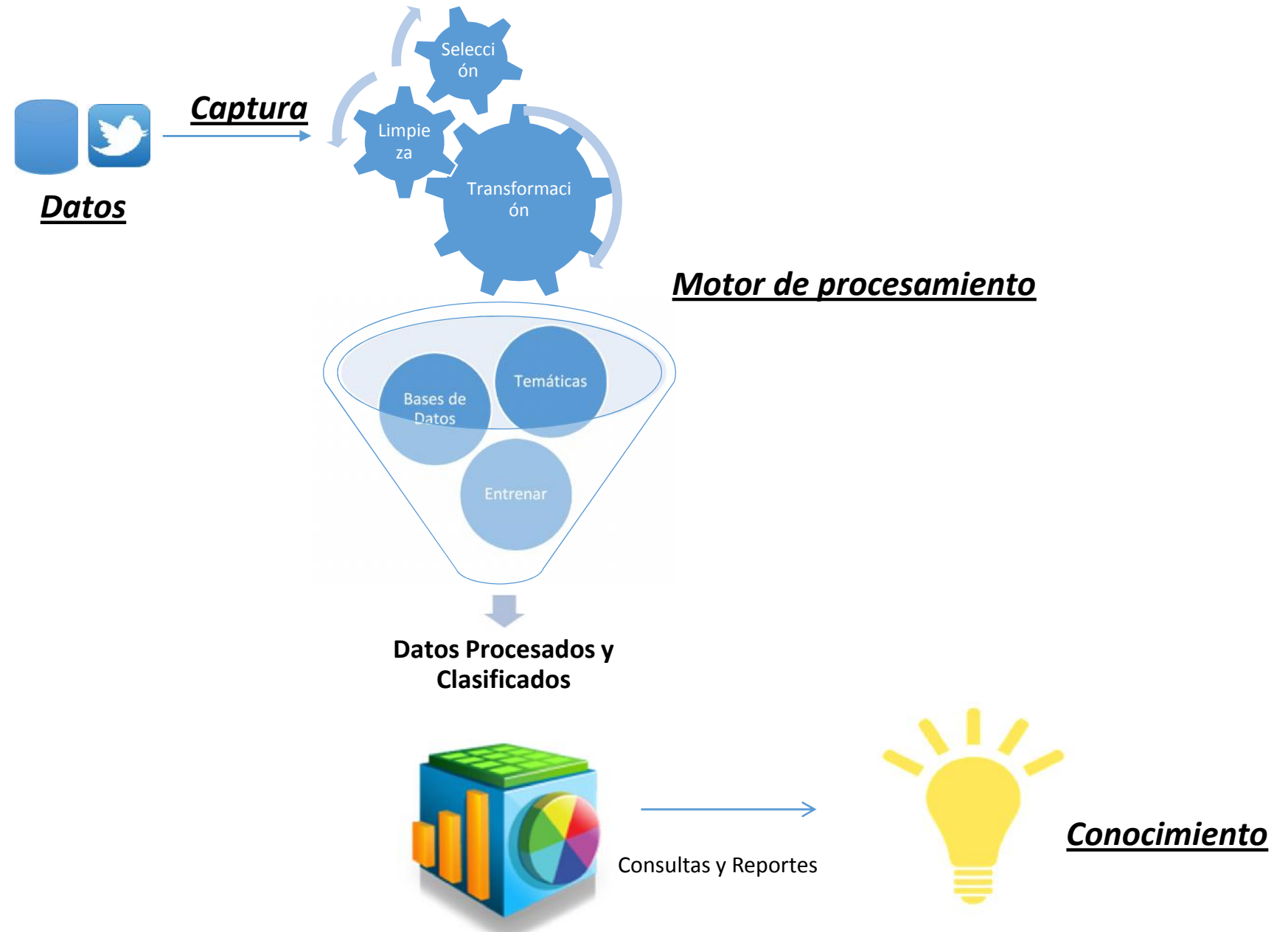
En esta etapa utilizamos el sistema VicData Analytics para el análisis de los datos y la presentación de resultados.





"La revolución de los datos"

# Procesamiento Big Data





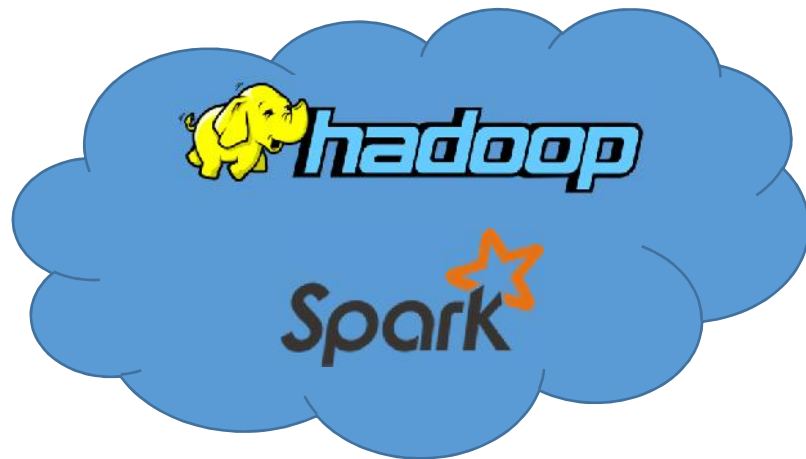
“La revolución de los datos”

# Herramientas para análisis BI



- Generación de Cubos de Datos
- Visualización de Resultados en gráficos de barras, pie, línea o mapas.
- Gráficos interactivos.
- Cuadros de mando a nivel operativo, táctico y estratégico.
- Análisis de sentimientos a los mensajes enviados por twitter en tiempo real.

# Plataforma en el Cloud



*Cloud Computing*

- Las herramientas tecnológicas para almacenamiento, procesamiento y análisis se encuentran alojadas en el Cloud de Google, lo cual nos permite disponer de un ahorro significativo de costos y agilidad en el escalamiento de infraestructura tecnológica.
  - *Almacén de Datos Hive*
  - *Spark*
  - *VicData Analytics*





**Demostración...**



**VicData**

“La revolución de los datos”

# ¿Preguntas?





**VicData**

“La revolución de los datos”

# Gracias.

